

Case Study: eduSource and the Athabasca University Digital Library Project

By Rory McGreal, Terry Anderson, Chris Hubick, Fuhua Lin, Mike Sosteric, Tony Tin, Oksana Yasynska

*Note: This is the author's original submitted article and has not been edited for style and content by ASTD.

Athabasca University has made the commitment to put all of its courses online as part of its Strategic University Plan. In pursuit of this goal, it has participated in the eduSource project, a pan-Canadian effort to build the infrastructure for an interoperable network of learning object repositories based on and incorporating international standards and specifications. At AU, this initiative has brought together professionals, academics, and other researchers into a team that has synchronized the efforts of different university centers in the creation of a common enterprise-wide university repository for learning objects. The project participants have created open source applications and intelligent agents as well as a suite of software tools that complement those of the other eduSource partners.

Athabasca University, Canada's Open University (AU), has been a partner in the pan-Canadian eduSource project, sponsored by CANARIE, Canada's advanced Internet development organization. AU, along with its five eduSource Canada partners has built the components and best practice guidelines to develop the infrastructure for a pan-Canadian testbed of linked and interoperable learning object repositories, based on national and international standards.

AU has produced an open source suite of tools for creating learning object repositories. The applications are based on and incorporate international standards and specifications and include a learning object repository (LOR) input application called ADLib, a digital reading room (DRR), an application for converting library records in MARC format to IEEE LOM records, which is the international standard for learning object metadata. There is an intelligent agent for enabling learners to quickly access answers to questions and other agents check for broken Web links and make recommendations to users. The software suite of tools includes Java bindings and interfaces to the LOR as well as identification and validation services.

The Digital Reading Room

The AU Digital Reading Room (DRR) began as an online repository of reserve articles and other course materials that were tagged for use in specific courses. Development of the DRR needed to meet the administrative and pedagogical needs of the AU users. Administrative users required a stable but flexible system, capable of incorporating different types of learning object resources into the DRR. Reliable authentication was considered essential, but ease of use, including a user-friendly method of data entry, modification, and retrieval was considered to be paramount by faculty. Student users also required stability and assured, reliable access to multiple types of learning resources at any time of the day or night, along with the capacity to connect, view and/or manipulate search results. The DRR was originally based on an SQL database, running on a Linux server. In March 2003, there were 40 Digital

Reading Files (DRF) with 2,100 learning objects. By February 2004, there were 80 DRFs with 4,100 learning objects.

The eduSource Canada Project brought to light a deficiency in the original DRR implementation. Namely, it was not interoperable with other repositories as it was not based on international standards. The implementers recognized the importance of rendering it interoperable using the IEEE LOM standard (IEEE, 2003) based on the CanCore metadata implementation profile (Athabasca University, 2004). At this time, further development of the DRR, other than essential operations, was suspended and work on an interoperable repository commenced. The DRR is being integrated into ADLib (Athabasca University Digital Library), which is presently based on a PostgreSQL database.

For more information on the DRR visit <http://library.athabascau.ca/drr>.

ADLib

AU has developed an IEEE LOM/CanCore compliant metadata repository application called ADLib. [ADLib](#) is a Web application for creating and storing standards compliant metadata records, and for storing their corresponding learning objects in the repository. Full courses, units of courses, tutorials, and articles are all learning objects at different levels of granularity that can be stored in the repository. The learning object metadata is used for searching and accessing learning objects. Metadata is simply a description of the learning object and it contains information such as title, description, and keywords. The advantage of this approach is that once the learning object has metadata, they become searchable. The ADLib repository can also be used by authors to keep track of their own material (e.g., *My Metadata* and *My Objects* tools).

ADLib facilitates the implementation of (and access to) single metadata records and learning objects, but it also enables users to group records or objects or create relationships between different objects. For example, a group of objects can be linked together to form a lesson which are linked together to form a course. The course can then be related to a particular programme. There are parallel relations among different LOM datasets or between the LOM records and the learning objects to which they refer.

ADLib also supports different modes, one for experienced users who are familiar with the IEEE LOM standard and another for novices. There are also tools for supporting different vocabularies. For example, ADLib provides the optional use of some AU-specific vocabularies along with IEEE LOM standard vocabularies. Vocabularies and modes can be easily set by user.

For more information about ADLib visit <http://edusource.athabascau.ca/ADLib.html>.

A MARC-IEEE LOM/CanCore Converter

The MARC – IEEE LOM/CanCore converter enables XML-based digital repositories to interact with and harvest metadata from MARC (Machine Readable Cataloguing) records that are used in libraries. MARC records contain hundreds of fields and were developed to describe print materials. This converter enables the simplification of the

records and their utilization in e-Learning. It will make re-purposed knowledge objects readily available alongside specially purposed learning objects.

The Converter is a Web-based program written in Java that takes the metadata of library materials in MARC format and converts them into IEEE Learning Object Metadata (LOM) format, displays a copy of the results on the user's screen, and saves the result to a database. The converter can be called a MARC to IEEE LOM/CanCore converter because it outputs data in IEEE LOM format refined by using CanCore recommendations, definitions and semantics.

The semantic mapping involved recording a path from each element in the source metadata standard to a semantically equivalent element in the target metadata standard. This process required in-depth knowledge and specialized expertise in the associated metadata standards. The mapping is an iterative process involving ongoing refinement, revision, and, even rethinking of element matches.

The converter is available at
<http://emd.athabascau.ca/courses/crosstalk/converter.html>.

FAQ Intelligent Agent

A FAQ agent has been developed to aid users in asking frequently asked questions (FAQs) about eduSource, learning objects, the Digital Reading Room/ADLib and the Library.

FAQ agent users have the option of

- searching by keywords, phrases
- asking a question
- searching using natural language
- browsing a list of questions in the database and linking directly to the answer
- browsing the list of keywords in the database and linking to questions and answers containing these keywords.

The agent provides an online interface for administrators, the Webmaster, or course designer to manage (add, delete, classify, and modify) the content of the FAQ database storing question-answer pairs. The success of a FAQ agent is dependent on a FAQ knowledge management system. The FAQ knowledge is organized in a hierarchy for quick access. We took an AU-centric approach to this challenge by creating state-of-the-art technology and augmenting it with an effective knowledge management process. The team combined the efforts of highly trained knowledge engineers with those of content experts who reviewed unanswered questions and worked to create new content to address these questions. They also adjusted the system configuration to improve search results. Because the FAQ agent is integrated with other AU and external repositories, the knowledge engineers were able to mine various logs for valuable content.

The FAQ agent is available at
<http://ADLibx.athabascau.ca/faqagent/jsp/kmanagement/faq.jsp>.

Broken Link Checking Agent

The broken link checking agent consists of a spider-like checking agent and an online interface. The checking agent can monitor all the links on the page of a specific URL. From the interface, users can cut and paste or enter a full URL that is of interest. Clicking on the "Submit" button activates it. If the agent detects any broken links in the Webpage of the URL, the URLs of the broken links will be displayed on the interface.

This agent is available at <http://216.123.230.77:8080/examples/servlets/brokenlink.html>.

Learning Object Recommendation Agent

This agent is designed to support the user in identifying and accessing learning objects according to personalized specifications that have been dynamically interpreted. The agent works on behalf of the user, monitoring the new arrival of learning objects in the learning object repository, notifying the user when relevant learning objects are deposited in the repository.

This agent is available at <http://ADLibx.athabascau.ca/lora/jsp/lora/index.jsp>.

LOM Java Binding

This is a Java API (Application Program Interface) binding of the IEEE LOM specification. This package defines an implementation neutral LOM Java interface (not classes) for exposing data objects corresponding to those elements defined by the LOM standard. This API can be used by Java programmers for representation of LOM data objects within their own software, or as an interoperable way to communicate LOM data objects with external third party software components. This API allows for the sharing of Java code between all those needing to process LOM data. The package also includes an Identifier interface, and a simple Learning Object interface.

This tool is available at <http://ADLib.athabascau.ca/~hubick/LOM>.

LOM Java Binding - Implementation

This is an implementation of the LOM Java Binding. This package contains: a LOMImpl class implementing those interfaces defined by the LOM Java Binding. The LOMImpl class includes functions designed to facilitate programmers writing code for instantiating the LOM objects it implements (i.e., building a LOM from a database result set). This implementation is designed to be extensible (by overloading a few methods) for use by those with additional non-standard LOM fields.

There are JAXB implementations for marshalling and unmarshalling LOM objects to and from XML. There are both JAXB LOM-XML implementation and Dublin Core implementations. The LOM-XML Unmarshalling code uses the VCard4J parser to process vCard entity data from LOM-XML instances, and the Marshaller has an option to expose this vCard data in a proprietary (VCard4J) XML format—possibly for easy use by XSLT. LOM date fields can be Marshalled in the standard LOM ISO-8660 format, or optionally in the RFC-1123 format used by HTTP and RSS (allowing XSLT directly on LOM-XML output without requiring date conversion)

This tool is available at <http://ADLib.athabascau.ca/~hubick/LOMImpl>.

LOR Java Interface

This package defines an implementation neutral Java API for a Learning Object Repository (LOR). The aim of this API is to provide a minimal set of standard repository functions from which a shared layer of services can be built up around any implementation. This package includes interface groups for working with both the metadata and learning object data itself. For example, there are metadata interfaces for retrieval; harvesting; storing; and searching; as well as a repository lifecycle control interface and a repository event listening interface. There is also a factory class which clients can use to transfer a reference to a LOR whose implementation can be configured at runtime.

This tool is available at <http://ADLib.athabascau.ca/~hubick/LOR>.

LOR Java Interface – Implementation

This is an abstract base implementation of the LOR Java Interface. This package provides the following:

- A LORImpl base class can be derived when creating a concrete LOR implementation. The base class provides data structures implementing result interfaces, as well as functionality for configuration loading, listener management, etc.
- A servlet provides access to most LOR interfaces. This includes direct access, form processing, HTTP basic authentication, and code for caching and applying XSLT to XML responses on the server side.
- A custom SOAP communications layer provides access to a LOR. This includes both a server side service implementation using Apache Axis, as well as an associated client that also implements the LOR interface, allowing for network-transparent access to a LOR.
- An OAI-PMH implementation is available, which is capable of supplying the standard Dublin Core metadata and LOM-XML.
- A GUI administration tool built with Swing is useful not only for administration of a repository, but also very helpful for testing and debugging during development.
- Other utilities are available for use when building a specific LOR implementation.
- On request, a sample XSLT I used for generating RSS feeds of recent repository objects (using XML from the servlet) can be supplied.

This tool is available at <http://ADLib.athabascau.ca/~hubick/LORImpl>.

LOR - LDAP Implementation

This is a partial implementation of the LOR Java API's using an LDAP server. This is achieved using a provided LDAP binding of the IEEE LTSC LOM schema. This package includes LOM_LDAP functionality for converting LOM objects to and from LDAP attributes. This implementation is fully LOM compliant, and capable of storing any valid LOM document in its entirety. Though, at this time, only a limited subset of the full LOR harvesting and searching capability is provided. It is expected that a common use of this package will be for identifier resolution.

<http://ADLib.athabascau.ca/~hubick/LOR-LDAP>

Identifier Resolution Service

This package provides a distributed service for resolving globally unique identifiers into their locations, using LDAP. Resolution records (identifier to location mappings) are stored in the LDAP server using LOM format data. The services for this system are simply an appropriate configuration of the software provided by the LOR-LDAP package and its dependencies.

This tool is available at <http://ADLib.athabascau.ca/~hubick/Resolver>.

LOM Validation Service

This service can be used to validate IEEE LTSC LOM XML format metadata against the LOM-XML Schema. The vCard data will also be run through a vCard (VCard4J) parser.

This tool is available at <http://ADLibx.athabascau.ca/LOMValidator>.

Globally Unique Identifiers

Important work has also been completed on unique identifiers. On the Web, most documents are referenced (and linked to) by the URL representing their network location. This works reasonably well when a document only has a single location, but naturally leads to a number of problems when copies of a document are made available from multiple locations. A major objective of the eduSource project was to create a system of sharing and reuse among documents, making a single location for a document, and thus identification using location as the search criterion is not possible. This problem can be solved by assigning each document a globally unique label. Documents are then free to automatically replicate throughout the network, and links can be created using identifiers.

However, once you have a location independent globally unique identifier, you still ultimately need some location from which to actually retrieve that document. Given an identifier, a resolution service is needed to determine the location from which a client can retrieve a desired document. What factors are relevant to the design and implementation of a resolution service? Any robust solution needs the ability to scale globally. This generally means that, like the Internet itself, it needs to be distributed. In the long run, you can't have one server resolving identifiers for everyone. A distributed system means you have no single point of failure, which creates fault tolerance, as well as scalability by dividing out the workload. A suitable architecture also needs to be open, it shouldn't force users into using any single implementation or proprietary or commercial technology. And the service needs to be available, or suitable for implementation, on a variety of hardware and software platforms. For these reasons the service should be architected using existing standards wherever possible.

Summary

These open source applications, agents, and tools have been developed by AU as part of its obligations in the eduSource project and as part of its commitment to put all of its courses online as part of the Strategic University Plan. These are designed not only for the implementation of a learning object repository infrastructure at AU, but also to ensure the interoperability of AU's repository with other Canadian and international LOR initiatives. At AU, this initiative has been used to create the tools necessary for implementation as well as bringing together professionals, academics, and other researchers into a team that has synchronized the efforts of different university centers in the creation of a common enterprise wide university repository for learning objects.

Rory McGreal is associate vice president of research at Athabasca University. **Terry Anderson** is a professor and Canada Research Chair in Distance Education at Athabasca University. **Tony Tin** is the electronic resources librarian at Athabasca University Library. **Chris Hubick** is an *eduSource SGML/Metadata Expert* for Athabasca University. **Fuhua Lin** is an associate professor of computing and information systems at Athabasca University. **Mike Sosteric** and **Oksana Yasynska** work on the *DEPTH Research/Database* at Athabasca University.

Additional Reading

- Athabasca University. (2004). *CanCore: Canadian core learning resource Metadata application profile*, Retrieved May 13, 2004, from <http://www.cancore.ca/>
- CANARIE. (2004). *CANARIE Home Page*, Retrieved May 13, 2004, from <http://www.canarie.ca>
- DCMI. (n.d.). *Dublin Core Metadata Initiative*, Retrieved May 13, 2004, from <http://dublincore.org/>
- Free Software Foundation, I. (2000, November 7). *GNU Lesser General Public License*, Retrieved May 19, 2004, from <http://www.gnu.org/copyleft/lesser.html>
- Hatala, M., Richards, G., Eap, T., & Willms, J. (2004). *The interoperability of learning object repositories and services: standards, implementations and lessons learned*, Retrieved May 4, 2004, from <http://www.sfu.ca/~mhatala/pubs/www04-interop-submit.pdf>
- Heaton, J. (2002). *Programming Spiders, Bots and Aggregators in Java*. Berkeley, CA: Sybex.
- Hubick, C. (2003). *Addressing: Identification and resolution of learning objects and metadata*, Retrieved May 3, 2004, from <http://adlib.athabascau.ca/~hubick/Resolver/addressing.html>
- IEEE. (2003). *Learning Technologies Standards Committee WG12: Learning Object Metadata*, Retrieved May 1, 2004, from <http://ltsc.ieee.org/wg12/>
- Lin, F. (2004). *Software agents for eduSource learning object repository*, Retrieved May 3, 2004, from http://edusource.athabascau.ca/files/agents-oscar_2.doc
- McGreal, R. (2004, May). Learning objects: A practical definition. *International Journal of Instructional Technology and Distance Learning (IJITDL)*, 1(5). Retrieved May 29, 2004, from <http://itdl.org/journal/>
- MIT. (2004). *Open Archives Initiative*, Retrieved May 13, 2004, from <http://www.openarchives.org/>
- Open LDAP Foundation. (2003, December 14, 2003). *Open LDAP*, Retrieved May 2, 2004, 2004, from <http://www.openldap.org/>
- Richards, G., Hatala, M., & McGreal, R. (in press). POOL, POND and SPLASH: Portals for online objects for learning. In R. McGreal (Ed.), *Online Education Using Learning Objects*. London: Routledge/Falmer.
- Scharfe, S. (2004). *Report on library applications of learning objects and ADLiB*, Retrieved May 3, 2004, from http://edusource.athabascau.ca/files/AU_library_LO_ADLiB.doc
- Yang Cao, Lin, F., McGreal, R., Anderson, T., Schafer, S., Friesen, N., et al. (2004, May 17-20) *Facilitating e-learning with a MARC to IEEE LOM metadata crosswalk application*. The 17th International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems. Ottawa: EA-AIE.